# ANA 515 Assignment 2

Hao Li

6/19/2022

# This dataset is called ADAE, which is a clinical adverse events dataset. This dataset is designed to collect the information of subjects who have adverse events during the whole clincal study. Moreover, the dataset is followed by CDISC standards (Clinical Data Interchange Standards Consortium).As for research questions, I want to know that what percentage of adverse events are caused by our study medication. The dataset is saved as XPT. format and it is fixed width if it is in a flat file. If it is binary, I would use Linux to open it.

###section2

```
library(foreign)
XPT<-read.xport("C:/Users/12052/Downloads/adae.XPT")
write.csv(XPT,"C:/Users/12052/Downloads/adae.csv")

adae<-read.csv("C:/Users/12052/Downloads/adae.csv")
```

#I selected read function to make my dataset readable by R, after transferring its format from XPT to CSV, then I gave the dataset a name as ADAE. the Package I used is foreign, which can help users easily transfer datasets in different format to be readable.

###section3

```
adae2<-adae %>%
subset(AGE>85)%>%
select(USUBJID, TRTA, SEX, AGE, RACE, TRTSDT,TRTEDT, AESEV)%>%
rename(gender=SEX)%>%
filter(gender=='M')%>%
print ()
```

```
##        USUBJID                TRTA gender AGE  RACE TRTSDT TRTEDT    AESEV
## 1  01-704-1241 Xanomeline High Dose      M  86 WHITE  19595  19640 MODERATE
## 2  01-704-1241 Xanomeline High Dose      M  86 WHITE  19595  19640     MILD
## 3  01-705-1199  Xanomeline Low Dose      M  87 WHITE  19617  19629 MODERATE
## 4  01-705-1199  Xanomeline Low Dose      M  87 WHITE  19617  19629 MODERATE
## 5  01-709-1285  Xanomeline Low Dose      M  87 WHITE  19441  19501     MILD
## 6  01-709-1285  Xanomeline Low Dose      M  87 WHITE  19441  19501     MILD
## 7  01-709-1285  Xanomeline Low Dose      M  87 WHITE  19441  19501     MILD
## 8  01-709-1285  Xanomeline Low Dose      M  87 WHITE  19441  19501     MILD
## 9  01-709-1285  Xanomeline Low Dose      M  87 WHITE  19441  19501     MILD
## 10 01-709-1285  Xanomeline Low Dose      M  87 WHITE  19441  19501     MILD
## 11 01-710-1002  Xanomeline Low Dose      M  88 WHITE  19737  19741     MILD
## 12 01-710-1002  Xanomeline Low Dose      M  88 WHITE  19737  19741     MILD
## 13 01-718-1328 Xanomeline High Dose      M  86 WHITE  19390  19466     MILD
## 14 01-718-1328 Xanomeline High Dose      M  86 WHITE  19390  19466     MILD
## 15 01-718-1328 Xanomeline High Dose      M  86 WHITE  19390  19466     MILD
## 16 01-718-1328 Xanomeline High Dose      M  86 WHITE  19390  19466 MODERATE
```

###section4

```
n_1 <-nrow(adae2)
n_2 <-ncol(adae2)
```

This dataframe has 16 rows and 8 columns. The names of the columns and a brief description of each are in the table below:

```
COLUMN<-c('USUBJID','TRTA','GENDER','AGE','RACE','TRTSDT','TRTEDT','AESEV')
DESCRIPTION<-c("SUBJECT ID NUMBER","ADMINISTRATE MEDICATION","GENDER","AGE","ETHNIC","TREATMENT
 START DATE","TREATMENT END DATE",
               "ADVERSE EVENTS SEVERITY LEVEL")
TABLE<-data.frame(COLUMN,DESCRIPTION)
knitr::kable(TABLE, COL.NAME=c("COLUMN","DESCRIPTION"))
```

| COLUMN | DESCRIPTION |
|---|---|
| USUBJID | SUBJECT ID NUMBER |
| TRTA | ADMINISTRATE MEDICATION |
| GENDER | GENDER |
| AGE | AGE |
| RACE | ETHNIC |

| COLUMN | DESCRIPTION |
|---|---|
| TRTSDT | TREATMENT START DATE |
| TRTEDT | TREATMENT END DATE |
| AESEV | ADVERSE EVENTS SEVERITY LEVEL |

###section5

```
adae3<-adae2 %>%
select(AGE,TRTSDT,TRTEDT)
summary (adae3)
```

```
##       AGE            TRTSDT          TRTEDT
##  Min.   :86.00   Min.   :19390   Min.   :19466
##  1st Qu.:86.00   1st Qu.:19428   1st Qu.:19492
##  Median :87.00   Median :19441   Median :19501
##  Mean   :86.75   Mean   :19507   Mean   :19556
##  3rd Qu.:87.00   3rd Qu.:19601   3rd Qu.:19632
##  Max.   :88.00   Max.   :19737   Max.   :19741
```

```
# missing values
sapply(adae3, function(x) sum(is.na(x)))
```

```
##    AGE TRTSDT TRTEDT
##      0      0      0
```

###summary

```
sum1<-summary(adae3)
sum2<-sapply(adae3, function(x) sum(is.na(x)))
```